



Universidad de Valladolid

FACULTAD DE TRADUCCIÓN E INTERPRETACIÓN

Grado en Traducción e Interpretación

TRABAJO DE FIN DE GRADO

Aproximación al reconocimiento voz-texto EN/ES aplicado a documentos audiovisuales: un estudio de caso

Presentado por D.^a María Otaño Jiménez

Tutelado por la Dra. M.^a Teresa Ortego Antón

Soria, junio de 2023

A mi familia, en especial a mi abuela Asun.

*«Al final solo importará
el amor que hemos recibido
y el que hemos dado».*

Índice

RESUMEN.....	5
ABSTRACT.....	5
1. INTRODUCCIÓN	6
1.1. Justificación.....	6
1.2. Competencias	7
1.3. Objetivos	7
2. MARCO TEÓRICO	8
2.1. Reconocimiento de voz-texto.....	8
2.2. La traducción automática y los sistemas de Google.....	8
2.2.1. La traducción automática	8
2.2.2. Los sistemas de traducción automática de Google.....	10
2.3. Evaluación de los sistemas de traducción automática.....	11
2.4. La posesición	11
2.4.1. Tipos de posesición.....	12
2.5. El subtitulado	12
3.1. Hipótesis.....	14
3.2. La selección de la muestra de análisis.....	14
3.2.1. Los subtítulos automáticos de YouTube	15
3.3. Parámetros de análisis	15
3.4. Las dificultades de traducción en los textos agroalimentarios	17
3.5. Proceso de análisis	17
4. ANÁLISIS Y RESULTADOS	20
4.1. Análisis de los errores de traducción.....	20
4.2. Resultados por tipología de error	28
4.2.1. Errores de precisión.....	28
4.2.2. Errores de fluidez	30
5. CONCLUSIONES	33
6. BIBLIOGRAFÍA.....	34

Índice de ilustraciones

Ilustración 1. Captura de pantalla de los subtítulos (en formato TXT) en inglés y español descargados a partir del enlace del vídeo de YouTube.	18
Ilustración 2. Captura de la hoja de cálculo de Excel con los segmentos y su número de caracteres.....	18

Índice de tablas

Tabla 1. Tipología de errores según Ortiz (2016: 63-64) adaptada de MQM.	17
Tabla 2. Tabla de análisis de los errores de traducción: segmentos originales, traducción automática y posesición.....	27

Índice de gráficos

Gráfico 1. Tipología de errores	28
Gráfico 2. Clasificación de los errores de precisión.....	29
Gráfico 3. Clasificación de las traducciones incorrectas.....	30
Gráfico 4. Clasificación de los errores de fluidez.	30
Gráfico 5. Clasificación de los errores de tipografía.....	31
Gráfico 6. Clasificación de errores gramaticales.....	32

RESUMEN

El avance en el desarrollo de la inteligencia artificial y, especialmente, en el procesamiento del lenguaje natural ha propiciado que se produzcan mejoras en los sistemas de reconocimiento de voz-texto, así como de traducción automática. Las plataformas de contenido multimedia se han beneficiado de estos avances y los han aplicado a los subtítulos traducidos, muchos de ellos generados automáticamente. En este contexto, el presente Trabajo de Fin de Grado (TFG) pretende comprobar si los sistemas de reconocimiento de voz-texto y de traducción automática de YouTube (iCloud) son capaces de ofrecer un producto final de calidad similar a la humana a la hora de trasvasar lingüísticamente un vídeo de temática agroalimentaria en lengua inglesa subtulado automáticamente al español, a partir de la selección de una muestra en la que analizamos los errores más recurrentes utilizando una metodología basada en el modelo MQM adaptado por Ortiz (2016). Los resultados que se desprenden nos permitirán vislumbrar los errores típicos y, a su vez, establecer pautas para perfeccionar los sistemas de reconocimiento de voz-texto y de traducción automática.

Palabras clave: sistema de reconocimiento voz-texto, traducción automática, error, subtítulo, posesición.

ABSTRACT

The development of artificial intelligence, especially applied to natural language processing, has led to improvements in automatic speech recognition (ASR) and machine translation systems. Multimedia content platforms and translated subtitles, many of them automatically generated, have benefited from these advances. In this context, this research project aims to test whether YouTube (iCloud) ASR and machine translation systems can provide a final product with human-like quality when linguistically transferring an English-language food video automatically subtitled into Spanish, selecting a sample in which we analyse the most recurrent errors using a methodology based on the MQM model adapted by Ortiz (2016). The results obtained will allow us to glimpse the typical errors and, in turn, to establish guidelines for perfecting automatic speech recognition and machine translation systems.

Keywords: automatic speech recognition system, machine translation, error, subtitling, post-editing.

1. INTRODUCCIÓN

En los últimos años, el sector audiovisual ha experimentado una fuerte globalización que ha tenido un impacto significativo en diversos ámbitos, como el económico, social, cultural y educativo. Con la evolución de la tecnología, la llegada de internet y el auge de los medios de comunicación, han surgido multitud de plataformas de vídeo gratuitas, y ninguna se ha hecho tan omnipresente como YouTube (Sedeño Valdellós, 2011: 12). YouTube es una de las plataformas de contenido de audio y vídeo más solicitadas del mundo. A partir de 2009, comenzó a incorporar subtítulos automáticos en sus vídeos y, en 2019, amplió su sistema utilizando la inteligencia artificial (IA) y aprendizaje automático, a través del impulso de los subtítulos en directo de Google I/O 2019, una tecnología innovadora que aprovecha algoritmos de vanguardia para convertir el audio en texto, mostrándolo en la pantalla en tiempo real y que permitió lograr una mayor precisión y velocidad por lo que respecta a la traducción en otros idiomas.

En la actualidad, ver películas, series y vídeos con subtítulos se ha convertido en un modo eficaz y sencillo de aprender un nuevo idioma. Además, los subtítulos han permitido el acceso igualitario a estos contenidos, además de mejorar notablemente la accesibilidad audiovisual para personas con discapacidad auditiva. No obstante, conviene subrayar que no todos los contenidos audiovisuales proporcionan subtítulos elaborados por traductores o subtituladores profesionales. En este sentido, las herramientas de reconocimiento de voz-texto han permitido desarrollar subtítulos y transcripciones automáticas que a su vez son traducidos por sistemas de traducción automática neuronal, lo que facilita la difusión de contenidos en diferentes idiomas.

El presente Trabajo de Fin de Grado (TFG), titulado *Aproximación al reconocimiento voz-texto EN/ES orientado a documentos audiovisuales: un estudio de caso*, pretende comprobar si los sistemas de reconocimiento de voz-texto y de traducción automática de YouTube (iCloud) son capaces de ofrecer un producto final de calidad similar a la humana a la hora de trasvasar lingüísticamente un vídeo de temática agroalimentaria en lengua inglesa subtitulado automáticamente al español, a partir de la selección de una muestra en la que analizamos los errores más recurrentes utilizando una metodología basada en el modelo MQM adaptado por Ortiz (2016). El mencionado trabajo se fundamenta en conceptos como el reconocimiento de voz-texto y la traducción automática, a lo que se añaden herramientas de traducción asistida por ordenador (TAO), en concreto, herramientas de subtitulado y otras herramientas informáticas.

1.1. Justificación

Los rápidos avances en el campo de la inteligencia artificial han dado lugar a notables mejoras en los sistemas de reconocimiento de voz-texto y de traducción automática, que están entrenados con sistemas de procesamiento lingüístico profundo y *machine learning*. Sin embargo, nuestra experiencia como usuarios ha revelado que los sistemas de Google producen errores y, por ende, estos se reflejan en la traducción automática de vídeos de YouTube. Así pues, existe una necesidad de mejorar diversos aspectos para lograr una mayor precisión en las traducciones. Esto requeriría el perfeccionamiento de la calidad del audio y del texto, así como los modelos de lenguaje y la capacidad de tener en cuenta el contexto para lograr traducciones más precisas.

En consecuencia, la tipología de errores que se producen en la traducción automática de subtítulos de YouTube por los sistemas de Google (iCloud) debe analizarse exhaustivamente para determinar los errores más recurrentes y sus causas subyacentes. En general, a pesar de que los avances tecnológicos han propiciado mejoras significativas en la traducción automática, todavía existen obstáculos que deben ser abordados para obtener una traducción precisa, fiable y similar a la humana.

La motivación que nos ha llevado a realizar una investigación en el ámbito de los sistemas de reconocimiento de voz-texto aplicados a la traducción automática surge de nuestro interés por las nuevas tecnologías y su implementación en el sector de la traducción profesional y de la traducción audiovisual. A lo largo del Grado en Traducción e Interpretación hemos tenido la oportunidad de cursar asignaturas como TIC para la Traducción, Informática Aplicada a la Traducción, Traducción Asistida por Ordenador (TAO) o Localización, que amplían nuestro conocimiento sobre los diferentes recursos informáticos, pero consideramos que aún no hemos aprovechado plenamente todo el potencial que estos pueden ofrecer. A esto se une el hecho de que el Consejo Social de la Universidad de Valladolid nos ha concedido por concurrencia competitiva una ayuda para colaborar en tareas de investigación en el Departamento de Lengua Española, bajo la tutela de Dra. Ortego Antón. En consecuencia, nos hemos propuesto profundizar en este campo iniciándonos en la investigación con este TFG y pretendemos alcanzar una serie de competencias y objetivos.

1.2. Competencias

En este Trabajo de Fin de Grado aplicamos las habilidades y conocimientos que hemos adquirido durante nuestra formación en el Grado de Traducción e Interpretación, junto con una serie de competencias específicas que se detallan en la guía docente de la asignatura Trabajo Fin de Grado del curso 2022-2023. Por lo que respecta a las competencias generales son: G1, G2, G3, G4, G5 y G6. Asimismo, se han desarrollado competencias específicas gracias a las diferentes asignaturas cursadas durante el Grado: E17, E18, E19, E21, E27, E34, E49, E50, E51.

1.3. Objetivos

El presente TFG tiene por objeto evaluar el potencial de los sistemas de reconocimiento de voz-texto y de la traducción automática de YouTube y determinar si pueden producir traducciones comparables a las realizadas por profesionales.

Este objetivo general, se concreta en los siguientes objetivos específicos:

- Determinar la importancia de los sistemas de reconocimiento de voz-texto y de la traducción automática por parte de YouTube, así como de los sistemas de Google.
- Detectar, corregir y profundizar en la tipología de errores cometidos por el sistema de traducción automática de los subtítulos de YouTube.
- Realizar una posesición de dicha traducción, teniendo en cuenta las restricciones que presenta el formato de los subtítulos.

Una vez establecidas las competencias y los objetivos de la investigación, profundizaremos en el marco teórico (Capítulo 2) para introducir los conceptos clave que vertebran el estudio. A continuación, detallaremos la hipótesis de partida y esbozaremos la metodología empleada (Capítulo 3) para estudiar la muestra de análisis (Capítulo 4) y, por último, presentaremos los resultados que nos permitirán extraer las conclusiones finales (Capítulo 5).

2. MARCO TEÓRICO

2.1. Reconocimiento de voz-texto

El reconocimiento de voz-texto hace referencia a un *software* que permite procesar la señal de voz humana o, en el caso del contenido audiovisual, previamente grabada y detectar la información lingüística para su transcripción a texto (Ávila Cornejo, 2016: 9). También se conoce como reconocimiento de habla. La entrada de habla se analiza mediante algoritmos de procesamiento del lenguaje natural y aprendizaje automático que identifican unidades léxicas en la entrada de voz y las convierten en una transcripción escrita.

Además, este *software* ofrece una traducción automática mediante el uso de modelos de aprendizaje automático basados en redes neuronales. Estos modelos se entrenan con grandes cantidades de datos textuales en varios idiomas, con el objetivo de que el sistema reconozca patrones y relaciones entre palabras y frases en diferentes lenguas. El modelo puede entonces utilizar esta información para proporcionar traducciones automáticas en tiempo real.

Los sistemas de reconocimiento de voz, ASR (*Automatic Speech Recognition*) pueden clasificarse en función de varios factores, como el modelo de locutor, el tamaño del vocabulario y la variabilidad del canal. Además, pueden ser independientes del hablante, dependientes del hablante o adaptativos al hablante. Aunque los sistemas de reconocimiento del habla independientes del locutor son más flexibles, su precisión es limitada en comparación con los sistemas ASR dependientes del locutor. Por el contrario, estos últimos ofrecen una mayor precisión, pero solo para hablantes previamente entrenados. Los sistemas adaptativos pueden aprender patrones de habla de nuevos hablantes. El tamaño del vocabulario es otro factor importante, ya que un vocabulario más amplio implica un sistema más complejo y una reducción en la precisión. El tamaño del vocabulario, a su vez, depende del idioma y el propósito del sistema (Malik *et al.*, 2020: 9421-9423).

En cuanto a su aplicación a la traducción, los sistemas de reconocimiento de voz-texto se utilizan para transcribir automáticamente el habla en diferentes idiomas y, a continuación, traducirla a otro idioma. Estos sistemas se basan en técnicas de procesamiento del lenguaje natural y aprendizaje automático para analizar el habla y generar una transcripción precisa. Una vez que se obtiene una transcripción fiable, el sistema aprovecha un traductor automático para traducir el texto a otro idioma.

Una vez definido este concepto, procedemos a discutir la traducción automática.

2.2. La traducción automática y los sistemas de Google

En este apartado delimitaremos la traducción automática, descubriremos los diferentes sistemas de TA, abordaremos su evaluación y profundizaremos en los sistemas de traducción automática que utiliza Google.

2.2.1. La traducción automática

La traducción automática (TA) se define como la producción de textos traducidos de una lengua natural a otra, con o sin intervención humana, y representa el cambio más profundo para la función del traductor (Cid-Leal *et al.*, 2019: 3).

El uso de sistemas de traducción automática es esencial para el desarrollo de las actividades de traducción automática. Los sistemas de traducción automática pueden clasificarse en cinco grupos en función de la tecnología subyacente: sistemas de traducción automática basados en reglas, basados en ejemplos, basados en estadística, híbridos o basados en redes neuronales (Sánchez Ramos y Rico Pérez, 2020: 575). Desarrollaremos brevemente los sistemas más interesantes para nuestra investigación.

2.2.1.1. Sistemas de traducción automática basados en reglas

La traducción automática basada en reglas, también conocida como *Rule Based Machine Translation* (RBMT), surgió en los años 70 (Forcada *et al.*, 2016: 151) y utiliza un conjunto de reglas gramaticales y léxicas predefinidas para traducir textos de un idioma a otro. Estas reglas se basan en la estructura gramatical y la sintaxis del idioma de origen y se utilizan para generar una traducción coherente y precisa en el idioma de destino. Para generar el texto en la lengua de llegada, el *software* realiza un análisis sintáctico del texto y construye una representación temporal. Este proceso se basa en léxicos extensos que contienen información morfológica, sintáctica y semántica, así como en amplios conjuntos de reglas.

No obstante, el mantenimiento de los diccionarios y reglas depende del trabajo humano constante, y el proceso de creación de un sistema puede ser costoso y largo. En general, la TA basada en reglas es menos precisa que los sistemas basados en aprendizaje automático, pero más fácil de implementar y requiere menos datos de entrenamiento (Maldonado y Liébana, 2021: 192).

2.2.1.2. Sistemas de traducción automática basados en estadística

La traducción automática estadística, método basado en corpus o *Statistical Machine Translation* (SMT) es un sistema de traducción que utiliza modelos estadísticos para traducir textos de un idioma a otro. Se basa en aprender de un corpus de textos de entrenamiento y utiliza técnicas de aprendizaje automático para asociar patrones en la lengua origen con su correspondiente traducción en la lengua meta. El proceso comienza con la recolección de datos y la preparación de estos mediante la selección y normalización de oraciones. El modelo de traducción determina conjuntos de posibles traducciones y sus probabilidades de ser correctas. La traducción final se produce a través de un proceso de búsqueda y elección de la traducción con mayor probabilidad, seguido por la corrección gramatical (Maldonado y Liébana, 2021: 193). La traducción automática estadística es precisa y rápida, pero puede tener problemas con idiomas o contextos fuera del corpus de entrenamiento y puede no capturar el tono, la sutileza y el significado implícito del texto original.

Los sistemas estadísticos, a diferencia de los modelos basados en reglas, no utilizan ningún tipo de información lingüística, sino que aprenden «automáticamente a traducir entre dos lenguas a partir de un corpus paralelo suficientemente grande de oraciones en la lengua original acompañadas de su traducción a la lengua meta» (Forcada *et al.*, 2016: 177).

2.2.1.3. Sistemas de traducción automática basados en redes neuronales

Un sistema de traducción automática basado en redes neuronales es un tipo de sistema de inteligencia artificial que utiliza redes neuronales para traducir automáticamente un texto de un idioma a otro. Estos sistemas se entrenan con grandes cantidades de traducciones previas que sirven de ejemplo para aprender a traducir. Los sistemas de traducción automática basados en redes neuronales suelen ser más precisos y presentan mayor fluidez que los basados en reglas, ya que son capaces de aprender y adaptarse a las complejidades y variaciones del lenguaje (Maldonado y Liébana, 2021: 194).

Estos sistemas de traducción automática utilizan redes neuronales recurrentes (RNN, por sus siglas en inglés) para procesar secuencias de texto y mantener un contexto a lo largo de la traducción. Estos sistemas representan numéricamente palabras y frases mediante vectores y utilizan técnicas de inteligencia artificial, como el aprendizaje profundo, para mejorar continuamente sus resultados (Casacuberta y Peris, 2017: 68-69). Ejemplos populares incluyen Google Translate y DeepL.

Las redes neuronales son algoritmos de aprendizaje automático que se basan en la forma en que funciona el cerebro humano y están compuestas por capas de nodos (neuronas) que se conectan entre sí. En el caso de un sistema de traducción automática basado en redes neuronales, las entradas son oraciones o palabras en un idioma de origen y las salidas son

traducciones en el idioma de destino. A través del proceso de aprendizaje automático, las conexiones entre las neuronas se fortalecen o debilitan para producir una traducción más precisa.

La traducción automática con sistemas basados en redes neuronales ha mejorado mucho, pero aún no es tan precisa como la realizada por un traductor humano y puede tener errores en contextos específicos. Por eso, es recomendable realizar una posesición del texto traducido, especialmente en contextos formales o con jergas específicas.

2.2.2. Los sistemas de traducción automática de Google

La plataforma de Google ofrece una gran variedad de servicios de traducción automática, que utilizan técnicas de procesamiento del lenguaje natural y aprendizaje automático para analizar el texto de entrada y generar traducciones precisas en el idioma de destino.

La traducción automática de Google se basa principalmente en el modelo de traducción neuronal o modelo de redes neuronales, que es una técnica avanzada de aprendizaje automático que utiliza redes neuronales artificiales para generar traducciones. Este modelo se entrena con grandes cantidades de datos de traducción, como textos paralelos en diferentes idiomas y corpus de textos (Ubillús Sánchez, 2022: 17).

Google Translate traduce desde una pluralidad de idiomas, se actualiza constantemente para mejorar su precisión y cobertura de idiomas; tiene un modo de traducción en tiempo real, que permite traducir texto o audio mediante el uso de la cámara del dispositivo o el micrófono, así como traducciones de idiomas hablados de manera no verbal como señas y gestos.

También ofrece otros servicios de traducción automática para empresas a través de su plataforma Google Cloud Translation API, que permite a las empresas integrar la traducción automática en sus propias aplicaciones y servicios. Google Cloud es una plataforma de computación en la nube que ofrece una variedad de servicios, entre los que se incluyen almacenamiento de datos, bases de datos, aplicaciones y herramientas de inteligencia artificial. Los servicios de Google que nos interesan para este trabajo principalmente son:

- La API Speech-to-Text: Se trata de un servicio de reconocimiento de voz a texto proporcionado por Google Cloud, es decir, basado en la nube, que permite a los desarrolladores convertir voz en texto en sus aplicaciones y servicios, mediante algoritmos de procesamiento de lenguaje natural y aprendizaje automático para analizar la entrada de audio y generar una transcripción precisa del habla. Asimismo, admite una variedad de idiomas y dialectos, entre los que se incluyen el inglés, español, francés, alemán y chino mandarín. También acepta varios formatos de audio, como WAV, FLAC y OGG. La API tiene funciones avanzadas como la detección automática del idioma, la identificación de comas y puntos finales en las oraciones, y el reconocimiento de varios hablantes. Además, ofrece características para mejorar la accesibilidad, como la transcripción automática de idiomas hablados de manera no verbal como señas y gestos y se puede utilizar junto a otras API de Google como la de traducción automática para tener una solución completa de reconocimiento de voz-texto y traducción.
- La API Translation: Es un servicio de traducción automática de Google que permite traducir textos de un idioma a otro mediante una solicitud HTTP enviada a una API. El servicio admite varios idiomas y puede detectar automáticamente el idioma de origen, al mismo tiempo que permite especificar el formato de salida del texto traducido y proporciona opciones avanzadas para el manejo de la traducción, como la capacidad de especificar un contexto.
- Live Captions en Google I/O 2019: Utiliza tecnología de reconocimiento de voz y traducción automática en tiempo real. Esta tecnología convierte el audio en texto y lo

muestra en pantalla en tiempo real. Los Live Captions se implementaron utilizando herramientas de inteligencia artificial y aprendizaje automático para garantizar una alta precisión y rapidez en la traducción (García, 2021).

Una vez descritos los sistemas de traducción automática específicos de Google, procedemos a abordar la evaluación de estos.

2.3. Evaluación de los sistemas de traducción automática

El proceso de evaluación de la calidad de una traducción conlleva una serie de decisiones complejas, tales como la elección de métricas, la forma en que estas deben ser aplicadas, la cantidad de texto que se debe examinar y cómo se deben interpretar los resultados obtenidos. Este proceso resulta especialmente relevante en la evaluación de un sistema de traducción automática, ya que permite determinar su viabilidad y mejorar su rendimiento (Metola, 2022: 12). Resulta igualmente relevante considerar el propósito de la evaluación, que puede estar orientado a la comparación entre distintos sistemas de traducción automática, el contraste de diferentes versiones de un programa, el análisis de los resultados en función de aspectos lingüísticos o la determinación del desempeño de un sistema de traducción automática en un ámbito especializado (Sánchez Ramos y Rico Pérez, 2020: 576).

Sánchez Ramos y Rico Pérez (2020: 576) ofrecen una clasificación de la evaluación según sea manual o automática:

- El tipo de evaluación manual o humana la realizan profesionales monolingües o bilingües y se caracteriza por su subjetividad (Sánchez Ramos y Rico Pérez, 2020: 34). La intención es que la traducción se lea como si fuera un texto original. Para evaluar las traducciones automáticas suelen emplearse tres modelos ampliamente utilizados, a saber, SAESJ2450, TAUS y MQM. SAESJ2450 fue una de las primeras iniciativas utilizada para sistematizar la evaluación en el sector automotriz. TAUS se centra en unificar criterios al resolver discrepancias entre empresas (Sánchez Ramos y Rico Pérez, 2020: 35). MQM busca armonizar estándares y es el modelo utilizado en este trabajo, adaptado por Ortiz (2016: 63-64).
- La evaluación automática, por otro lado, compara una traducción de referencia con una traducción automática. Las principales métricas son BLEU, METEOR y TER. BLEU compara dos traducciones con un algoritmo, METEOR también compara dos textos y detecta equivalentes semánticos, y TER contabiliza la cantidad de ediciones necesarias para corresponder con un texto de referencia.

En resumen, hemos abordado la importancia de la evaluación de la calidad en la traducción automática y hemos explorado las diferentes metodologías utilizadas para evaluar la calidad de los sistemas de TA. Sin embargo, es importante destacar que en este trabajo no nos centraremos en los sistemas de evaluación automática, ya que no tenemos acceso a este tipo de sistemas, por lo que nos centraremos en la evaluación manual y emplearemos una adaptación propuesta por Ortiz (2016: 63) del modelo MQM.

2.4. La posesición

En el ámbito de la traducción automática, Wagner (1985: 1) fue uno de los primeros en abordar el concepto de posesición, al que definió como el proceso de corrección de un texto pretraducido, en lugar de la traducción completa «desde cero».

Otros autores como Veale y Way (1997: 12) describen la posesición como las acciones realizadas por lingüistas y editores para corregir los resultados de la traducción automática. Para Allen (2003: 297), la posesición consiste en ajustar, editar, modificar o corregir un texto pretraducido mediante sistemas de traducción automática desde una lengua fuente a una o más lenguas de llegada.

La definición más completa y referencial la encontramos en O'Brien (2011: 197): «*the correction of raw machine translated output by a human translator according to specific guidelines and quality criteria*». Para garantizar la calidad del texto final y su adecuación a las necesidades del cliente y al propósito del texto, es fundamental establecer directrices y guías de estilo.

De hecho, la finalidad última de la posesición es «mejorar el resultado obtenido por un sistema de traducción automática, que no tiene por qué ser necesariamente perfecto» (Sánchez Ramos y Rico Pérez, 2020: 78).

Una vez revisadas las diferentes definiciones propuestas para la posesición, es importante identificar sus distintos tipos existentes.

2.4.1. Tipos de posesición

Según Allen (2003: 297-318), existen dos tipos de posesición:

- a) Posedición completa o *full post-editing*, se enfoca en textos traducidos para su publicación, con una mayor calidad en el texto meta. Se realizan todos los cambios necesarios, pero no se recomienda modificar el estilo, ya que dicha modificación puede disminuir la productividad del sistema de TA. Este tipo de posesición se enfoca en garantizar la corrección gramatical, sintáctica y semántica de la traducción; comprobar la correcta traducción de la terminología; asegurar que no haya adiciones u omisiones; corregir el contenido ofensivo, inadecuado o culturalmente inaceptable y aplicar reglas de ortografía y puntuación.
- b) Posedición rápida o *light post-editing*, implica realizar cambios mínimos en el texto para facilitar la comprensión del lector. Esta técnica se utiliza cuando el objetivo principal es garantizar que el lector pueda entender la idea general del texto. No requiere cambios significativos, sino que se centra en mejorar la fluidez y claridad del texto traducido.

Una vez definidos los tipos de posesición, abordaremos el concepto de subtítulo y describiremos las dificultades y características de esta técnica de traducción.

2.5. El subtítulo

El subtítulo es una técnica de traducción audiovisual que permite a los espectadores entender el contenido audiovisual en una lengua distinta a su primera (Ávila *et al.*, 2020: 40). Consiste en añadir texto a una grabación de audio (como una película, vídeo o un programa de televisión) para proporcionar una traducción de diálogos o descripciones de audio en otro idioma o para facilitar la accesibilidad a personas con discapacidad auditiva. Los subtítulos aparecen en la parte inferior de la pantalla y, generalmente, se sincronizan con el audio para que las palabras aparezcan en el momento adecuado.

La cantidad de caracteres permitidos por línea y por segundo en la subtitulación puede variar dependiendo del estándar o requerimiento. Por lo general, se recomienda que una línea de subtítulo no supere los 45 caracteres y que los subtítulos se muestren durante al menos un segundo y un máximo de 7 segundos. De manera general, se utiliza la regla de los seis segundos (d'Ydewalle *et al.*, 1987; Brondeel, 1994; en Díaz Cintas, 2003). De acuerdo con este sistema, un espectador medio puede leer y comprender dos líneas con subtítulos de un total de 70 caracteres (35 por línea) en seis segundos (Rebollo Trigueros, 2018: 14).

Para que los subtítulos sean efectivos, deben ser fáciles de leer y no interferir en la experiencia del vídeo. En cuanto a la segmentación, los profesionales sugieren seguir reglas sintácticas y mantener las unidades lingüísticas juntas en una misma línea. Esto es importante porque «los

lectores no esperan hasta el final de la oración para interpretarla, sino que intentan encontrarle sentido mientras leen» (Szarkowska y Gerber, 2018: 2).

Al examinar una escena de un documento audiovisual como una imagen fija, con solo un vistazo el usuario podrá comprender el contenido. Sin embargo, debido a la rapidez y automatismo de la percepción, los usuarios no son plenamente conscientes del tiempo necesario para comprender lo que están viendo. Existen dos formas en las que se realiza la identificación de la escena y sus elementos. Por un lado, se pueden identificar los objetos y, a continuación, entender lo que se está viendo; por otro lado, se puede percibir el significado y diseño de la escena antes de identificar los objetos específicos (Díaz *et al.*, 2022: 16).

No obstante, la traducción no solo implica la transmisión del significado literal de las palabras, sino también la transmisión de la intención y las emociones que subyacen, por lo que extraer el sentimiento del hablante de flujos de audio naturales como en YouTube es todo un reto. Varios factores contribuyen a la dificultad de la tarea, a saber, el reconocimiento automático del habla (ASR) espontánea, entornos de fondo desconocidos, características variables de la fuente y el canal, acentos, etc. Ocurre lo mismo con el lenguaje no verbal, es decir, los movimientos, los gestos corporales o faciales de los protagonistas, todos ellos afectan al mensaje que se quiere transmitir en la grabación, por lo que se han de tener en cuenta a la hora de traducir o subtítular (Kaushik *et al.*, 2013: 239).

2.6. Recapitulación

En este capítulo hemos explorado y definido los conceptos clave relacionados con el objeto de estudio y hemos revisado la literatura relevante para establecer un marco teórico sólido y coherente, que nos permitirá fijar la metodología que aplicaremos a nuestro análisis y verificar o refutar la hipótesis que formulamos a continuación.

3. HIPÓTESIS Y METODOLOGÍA DE ANÁLISIS

En este capítulo, en primer lugar, perfilaremos la hipótesis que guía este trabajo. A continuación, elegiremos y justificaremos la muestra de análisis y explicaremos el proceso mediante el cual YouTube genera los subtítulos automáticos en sus vídeos. Posteriormente, acotaremos la muestra según su temática, la agroalimentación, y definiremos las características y dificultades asociadas a los textos de esta especialidad. Por último, especificaremos el modelo que vamos a utilizar para analizar los errores en la traducción automática y aportaremos una explicación del proceso para realizar el análisis.

3.1. Hipótesis

Los subtítulos generados por el sistema de reconocimiento automático de voz (ASR) y de traducción automática del inglés al español en YouTube contienen errores. Además, consideramos que la calidad es insuficiente para capturar la complejidad y diversidad de los términos, por lo que dichos subtítulos requieren de una posesición manual para asegurar la precisión en la transferencia interlingüística.

Para verificar o refutar esta hipótesis, seleccionamos una muestra de análisis representativa de los subtítulos generados por el sistema de reconocimiento automático de voz y la traducción automática de inglés a español de YouTube centrada en el sector agroalimentario. A continuación, realizaremos un análisis de los errores encontrados en los subtítulos automáticos para evaluar la calidad de la traducción. Para evaluar la precisión de los resultados, utilizaremos la métrica MQM, concretamente los parámetros de exactitud (*accuracy*) y fluidez (*fluency*).

3.2. La selección de la muestra de análisis

A pesar de que existen múltiples plataformas de contenido audiovisual —Vimeo, Vube, Dailymotion o Yahoo Screen, entre otras— YouTube, creada en 2005, se ha convertido en un fenómeno audiovisual sin precedentes en términos cuantitativos. La plataforma cuenta con más de mil millones de usuarios, lo que equivale a un tercio de todos los usuarios de internet. Cada minuto se suben a YouTube 100 horas de vídeo, y cada mes se visualizan en la plataforma más de 6000 millones de horas de vídeo, lo que supone casi una hora por cada habitante del planeta. Estos datos ponen de manifiesto la relevancia de YouTube en la vida de millones de personas y su impacto en la manera de producir, distribuir y consumir contenidos audiovisuales (Rendueles y Sádaba, 2019: 108).

YouTube es un medio abierto y colaborativo que permite a los usuarios de cualquier parte del mundo compartir sus producciones sin restricciones. Esta idea está vinculada directamente con la necesidad del medio de generar subtítulos en diferentes idiomas, para dar respuesta a la demanda global. La plataforma de YouTube ha democratizado la producción y distribución de contenido audiovisual, lo que ha llevado a una enorme cantidad de usuarios a crear contenido en diferentes idiomas y para distintas audiencias. Para garantizar que los vídeos sean accesibles para todos, se ofrecen subtítulos en varios idiomas para que los usuarios puedan comprender y disfrutar del contenido, fomentando así la comunicación y el intercambio cultural a nivel mundial. Por este motivo, YouTube cuenta con un sistema de reconocimiento de voz para generar automáticamente subtítulos en diferentes idiomas y su posterior traducción con los sistemas de traducción automática de Google (iCloud).

La traducción automática de los subtítulos es una herramienta muy útil para aquellos usuarios que no hablan el idioma original del vídeo, ya que les permite entender el contenido y comunicarse de manera más efectiva. Sin embargo, la calidad de la traducción automática puede variar significativamente dependiendo del idioma, la especialización del contenido y la complejidad semántica de los términos utilizados.

Para este trabajo, hemos escogido el vídeo *How to make paella* de Gordon Ramsay, un tutorial de cocina de 4 minutos y 23 segundos en el que el famoso chef británico explica a Juliet Oliver cómo preparar una «auténtica» paella valenciana. El contenido se presenta en inglés y se enfoca en la descripción detallada de los ingredientes necesarios, algunos de los cuales no son típicos de la receta original, así como en la técnica adecuada para cocinar paella. Durante el vídeo, Ramsay utiliza una variedad de ingredientes frescos y locales, y explica cada paso del proceso en un tono relajado y amigable.

La relación entre los participantes se corresponde con experto-iniciado (*expert to initiates*) (Ruiz Romero, 2020: 51). La situación comunicativa se establece para destinatarios con un nivel de conocimiento inferior al del emisor en el área de especialidad, es decir, Gordon Ramsay es un chef experto y se dirige a un público lego. Se utilizan términos especializados y, con frecuencia, se incluye el significado de los términos que se consideran complejos o desconocidos (Durán y del Moral, 2014: 48).

En el caso de los vídeos de carácter agroalimentario, la precisión en la traducción de los términos especializados es esencial para garantizar la correcta comprensión de la información por parte de la audiencia, así como para evitar errores y malentendidos en la comunicación intercultural.

3.2.1. Los subtítulos automáticos de YouTube

De acuerdo con Kaushik *et al.* (2013: 239) la tecnología de YouTube para generar subtítulos automáticos en vídeos se basa en técnicas de procesamiento del lenguaje natural (PLN). Dicha tecnología utiliza modelos de lenguaje basados en aprendizaje automático para analizar el contenido de audio de un vídeo y generar una transcripción textual y precisa del mismo.

Para generar los subtítulos, YouTube primero divide el audio en fragmentos y los analiza para identificar las palabras habladas. A continuación, utiliza un modelo de lenguaje para determinar la gramática y el contexto correctos y generar una transcripción en formato de texto. En este sentido, YouTube utiliza varios modelos de lenguaje para generar subtítulos automáticos y cada uno de ellos tiene diferentes niveles de precisión. Los subtítulos generados automáticamente son solo una primera aproximación, así que es común que contengan errores y omisiones o palabras mal escritas, de manera que la corrección y edición de los subtítulos automáticos es una tarea importante para mejorar la precisión.

Además, el lenguaje no verbal es fundamental a la hora de subtítular un vídeo, ya que proporciona información valiosa sobre el contexto, las emociones y el sentimiento de los personajes en pantalla. Los gestos, la postura, la mirada, entre otros aspectos no verbales, complementan y enriquecen el significado de las palabras habladas. Para detectar el sentimiento de los vídeos de YouTube, se utiliza un sistema que primero elimina el audio y, a continuación, utiliza la configuración de reconocimiento para obtener las mejores transcripciones de voz. Tras la decodificación, se analiza el texto con un etiquetador POS (*part-of-speech*) para obtener características de sentimiento textual y, por último, se utilizan modelos de sentimiento basados en ME (*Maximum Entropy*) para detectar la polaridad del sentimiento (Kaushik *et al.*, 2013: 240).

3.3. Parámetros de análisis

Para detectar y clasificar los errores utilizamos la métrica MQM (*Measurement of Quality in Machine Translation*), una métrica que permite evaluar la calidad tanto de las traducciones automáticas como de las traducciones humanas (Lommel *et al.*, 2014: 456). Esta métrica, mide la calidad en términos de coherencia, cohesión, precisión, corrección, fluidez, terminología y estilo. La métrica está basada en una lista de verificación de criterios de calidad y utiliza un sistema de puntuación para calificar la traducción en relación con los criterios establecidos.

MQM consta de tres fases:

- 1) una etapa preliminar para revisar, evaluar, preparar el texto de partida y el texto meta y determinar el número de palabras de la evaluación mediante una herramienta TAO (texto original: 706 palabras; texto meta: 592 palabras);
- 2) una etapa de anotación de errores en la que el evaluador, en este caso, el traductor, ha de examinar el texto traducido comparándolo con el texto original y las especificaciones del encargo y anotar, es decir, identificar, marcar y asignar el tipo de error de acuerdo con la métrica;
- 3) por último, se calcula la puntuación global de calidad (en nuestro caso, en este último paso, procederemos a clasificar los tipos de errores más recurrentes).

Dado que MQM es una métrica muy extensa, optamos por emplear el modelo simplificado desarrollado por Ortiz (2016: 63), en el que se utilizarán las categorías de análisis relativas a «*accuracy, issue, type and mechanical issues included in fluency*» (Lommel *et al.*, 2014), ya que se consideran las más relevantes para este trabajo. Los errores relacionados con la precisión se refieren a problemas de terminología, traducción incorrecta, omisiones, adiciones o contenido sin traducir. Por otro lado, los errores de fluidez se centran en aspectos como la ortografía, la tipografía, la gramática y la ininteligibilidad del texto.

A C C U R A C Y	Terminology	A term is translated with a term other than the one expected for the domain or otherwise specified.	
	Mistranslation	The target content does not accurately represent the source content.	
		Overly Literal	The translation is overly literal.
		False Friend	The translation has incorrectly used a word that is superficially similar to the source word.
		Should not have been translated	Text was translated that should have been left untranslated.
		Date/Time	Dates or times do not match between source and target.
		Unit Conversion	The target text has not converted numeric values as needed to adjust for different units.
		Number	Numbers are inconsistent between source and target.
		Entity	Names, places or other “named entities” do not match.
	Omission	Content is missing from the translation that is present in the source.	
	Addition	The target text includes text not present in the source.	
Untranslated	Content that should have been translated has been left untranslated.		
F L U E N C Y	Spelling	Issues related to spelling of words.	
		Capitalization	Issues related to capitalization.
		Diacritics	Issues related to the use of diacritics.
	Typography	Issues related to the mechanical presentation of text. The category should be used for any typographical errors other than spelling.	
		Punctuation	Punctuation is used incorrectly for the locale or style.
		Unpaired quote marks or brackets	One of a pair of quotes or brackets is missing from the text.
	Grammar	Issues related to the grammar or syntax of the text, other than spelling and orthography.	
		Morphology	There is a problem in the internal construction of a word.
		Agreement	A word is in the wrong part of speech.
		Word order	The word order is incorrect.
	Function words	A function word is used incorrectly.	

Unintelligible	The exact nature of the error cannot be determined. Indicates a major break down in fluency.
----------------	--

Tabla 1. Tipología de errores según Ortiz (2016: 63-64) adaptada de MQM.

Una vez explicada la métrica, procedemos a describir las dificultades de traducción típicas de los textos agroalimentarios.

3.4. Las dificultades de traducción en los textos agroalimentarios

La traducción agroalimentaria, como cualquier otro campo de especialidad, conlleva ciertos retos. Epstein (2018) reconoce cuatro tipos de dificultades principales y ofrece sus posibles soluciones.

- 1) La disponibilidad de los ingredientes. A pesar de la popularidad creciente de la cocina y la tendencia actual de utilizar ingredientes étnicos, no todos los productos están disponibles en todos los países y algunos solo se encuentran a precios exorbitantes. La sustitución podría ser una solución adecuada en algunos casos, pero no es suficiente que el traductor sustituya ingredientes al azar. El traductor debe ceñirse lo máximo posible al original y, si se ofrecen ideas para sustituir ingredientes, tiene que explicar el por qué.
- 2) Los cortes de la carne. Los cortes de carne no son necesariamente los mismos en los distintos países, así que en muchos casos será necesario preguntar a expertos y utilizar materiales de referencia.
- 3) Las medidas. Existen diferentes sistemas de medición en todo el mundo, lo que implica que no es suficiente con recurrir únicamente a páginas web de conversión. En los casos en que hay que cambiar las medidas, se puede optar por dos estrategias posibles. La primera es que el editor conserve las medidas y ofrezca una tabla de conversión al final del documento. No obstante, esta opción puede resultar tediosa para el receptor y su efectividad puede variar en función del tipo de documento en cuestión. Otra opción es utilizar nuevas medidas basadas en el sistema de la cultura de destino. Esto puede hacerse mediante sustitución completa (el traductor comprueba la receta y cambia las medidas asegurándose de que tengan sentido en el contexto) o sustitución y retención (implica tanto cambiar la receta al sistema de medida de la cultura meta como mantener la medida del original entre paréntesis).
- 4) Los utensilios, ollas y sartenes. Al igual que ocurre con los ingredientes, algunos países utilizan distintos utensilios, ollas, sartenes y otros instrumentos de cocina esenciales, o pueden emplear palabras radicalmente distintas para designar un utensilio similar. Se pueden solventar estas dificultades mediante la reescritura de la oración, la inclusión de una explicación sobre el objeto en cuestión, la búsqueda del término adecuado en glosarios para otros objetos, la adición de notas por parte de los traductores u otros materiales extratextuales necesarios.

Como podemos ver, existen diversas soluciones a las dificultades de traducción agroalimentaria. No obstante, la muestra de análisis seleccionada es un vídeo de carácter divulgativo, por lo que su grado de especialización es bajo, aunque podemos encontrar algunas dificultades descritas anteriormente, como pueden ser las medidas o los ingredientes.

A continuación, procedemos a describir el proceso que vamos a seguir durante el análisis.

3.5. Proceso de análisis

Durante el análisis, se ha seguido un proceso sistemático y riguroso que ha permitido la obtención de resultados precisos y fiables. En primer lugar, se ha procedido a la descarga de los subtítulos generados automáticamente por YouTube en inglés y español. A continuación, se ha llevado a cabo una comparación entre los subtítulos en inglés y el audio original del

vídeo, realizando las modificaciones necesarias para garantizar una correspondencia exacta entre ambos.

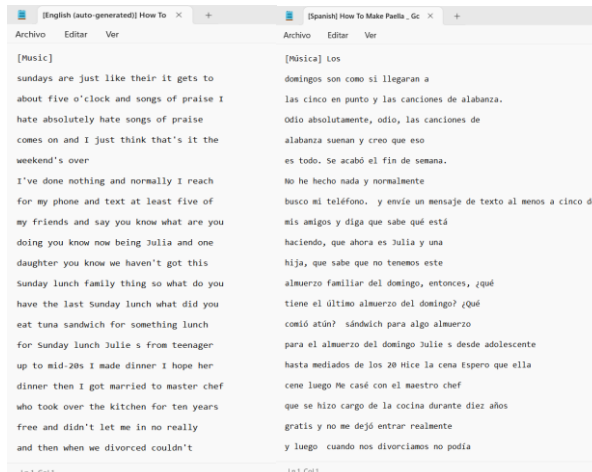


Ilustración 1. Captura de pantalla de los subtítulos (en formato TXT) en inglés y español descargados a partir del enlace del vídeo de YouTube.

Una vez completada esta, se ha creado una hoja de cálculo de Microsoft Excel con diversas columnas, tales como la segmentación original según la transcripción, el número de caracteres, el segmento traducido automáticamente, la posesición y el número de caracteres de esta última. Para proceder a la segmentación de los subtítulos automáticos, se ha tenido en cuenta el número de caracteres por línea, siendo recomendable que cada línea de subtítulo no contenga más de 42 a 45 caracteres.

Segmento original - Transcripción	núm caracteres	Subtítulo traducido automáticamente	Posedición	NÚM CARACTERES
1 Sundays are just a nightmare	29	Los domingos son como...	los domingos son una pesadilla	30
2 it gets to about five o'clock	29	si llegaran a las cinco en punto	llegan las cinco	16
4 and songs of praise. I hate,	28	y las canciones de alabanza, odio,	y las canciones de alabanza, odio,	34
5 absolutely hate songs of praise	31	absolutamente odio, las canciones de alabanza	odio profundamente, las canciones de alabanza	45
6 Comes on and I just think "that's it,	37	suenan y creo que eso es todo. Se acabó el fin	suenan y pienso: eso es todo,	31
7 the weekend's over, I've done nothing"	38	de semana. No he hecho nada	se acabó el fin de semana, no he hecho nada"	44
8 And normally I reach for my phone and	38	y normalmente busco teléfono, y	Y normalmente cojo mi teléfono y	33
9 text at least five of my friends,	33	envíe un mensaje de texto al menos a cinco de mis amigos	escribo al menos a cinco de mis amigos,	39
10 and say, you know, "what are you doing"	39	y diga que sabe qué está haciendo,	y les digo, ya sabes, "¿qué estás haciendo?"	43
11 you know, now being Juliet and one daughter	43	que ahora es Julia	sabes, ahora siendo Juliet	27
12 you know, we haven't got	24	que sabe que no tenemos este	sabes, no tenemos eso	21
13 this Sunday lunch family thing	31	almuerzo familiar del domingo,	de la comida familiar de los domingos	37
14 (Gordon) So what did you have	29	entonces, ¿qué tiene el último	Entonces, ¿qué tuvisteis	25
15 for last Sunday lunch?	22	almuerzo del domingo?	el domingo pasado?	18
16 What did you eat? (Girl) Tuna sandwich	38	¿Qué comió atún? sándwich	¿Qué comiste? Un sandwich de atún	35
17 For Sunday lunch?	17	para algo almuerzo	¿Para comer el domingo?	23
18 For Sunday lunch	16	para el almuerzo del domingo	Exacto.	7
19 From teenager up to mid-20s	27	desde adolescente hasta mediados de los 20	De adolescente hasta mediados de los 20	40
20 I made dinner, a proper dinner.	31	Hice la cena. Espero que ella cene	preparé la cena, una cena decente	33
21 Then I got married to a Master Chef	35	Me casé con el maestro chef	Luego me casé con un Master Chef	32
22 who took over the kitchen for ten years	38	que se hizo cargo de la cocina durante diez años	que se encargó de la cocina durante diez años	45
23 Really?	7	(omisión)	¿En serio?	10
24 And didn't let me in	20	y no me dejó entrar	Y no me dejó entrar	19
25 Oh really?	10	Realmente	¿En serio?	10
26 and then when we divorced	25	y luego cuando nos divorciamos	y cuando nos divorciamos	24
27 couldn't remember what to do.	29	no podía recordar qué hacer.	no podía recordar qué hacer.	28
28 So, your husband, your ex-husband	33	así que ve a tu esposo, a tu ex esposo	Así que, tu marido, tu ex-marido	32
29 To blame for this one. Absolutely.	34	no le gusta esto,	es el culpable. Desde luego.	28

Ilustración 2. Captura de la hoja de cálculo de Excel con los segmentos y su número de caracteres.

Posteriormente, se ha llevado a cabo una comparación entre los subtítulos en inglés y los de español, detectando, señalando y clasificando los diferentes tipos de errores encontrados mediante el uso de colores. Este proceso ha permitido una posesición manual de los subtítulos con el fin de obtener una mayor precisión y calidad en la traducción.

Finalmente, se ha procedido a la elaboración de gráficos con resultados cuantitativos (porcentajes) que han permitido identificar cuáles son los tipos de errores más recurrentes en los subtítulos traducidos automáticamente por YouTube.

Con el análisis de los errores más frecuentes en la traducción automática de YouTube, buscamos identificar los patrones y las causas de los errores, lo que permitirá mejorar la calidad de la traducción y de los sistemas de traducción automática y reconocimiento de voz-texto en el futuro. Este análisis también proporcionará información útil para los usuarios que utilizan este sistema de traducción automática, permitiéndoles comprender mejor los errores que pueden surgir en la traducción de subtítulos y, por ende, mejorar la comprensión de los contenidos en diferentes idiomas.

Una vez explicada la metodología que utilizaremos para el análisis, procedemos a aplicarla en la muestra seleccionada.

4. ANÁLISIS Y RESULTADOS

En este capítulo se presenta un análisis de los errores más frecuentes producidos por el sistema de YouTube en la traducción al español de los subtítulos de un vídeo en inglés, y utilizando la metodología descrita en el capítulo previo.

4.1. Análisis de los errores de traducción

En la Tabla 2 se presentan los segmentos originales (transcripción), los segmentos traducidos automáticamente y una posesición completa de los subtítulos del vídeo.

Es importante destacar que la tabla se ha realizado previamente en Excel, midiendo el número de caracteres por línea permitidos por los subtítulos a través de la fórmula (=LARGO) para poder realizar una división correcta de los segmentos (Ver Ilustración 2).

En primer lugar, en la columna de la izquierda, se encuentran los segmentos originales. Estos segmentos son transcripciones obtenidas de la descarga de los subtítulos generados automáticamente por YouTube en inglés que se han comparado con el audio original del vídeo y se han modificado de manera pertinente para asegurar una correspondencia exacta.

Como se revela en el [Anexo 1](#), la columna que muestra los subtítulos automáticos descargados presenta deficiencias en cuanto a la segmentación adecuada de acuerdo con los principios de subtitulación, dicha segmentación no logra mantener la cohesión de las unidades lingüísticas en una sola línea, lo que hace el texto casi indecifrible para el usuario. Para garantizar una transcripción de la máxima exactitud y rigor, se han introducido los cambios en negrita, entre los que destacan las mayúsculas y los signos de puntuación debidos a la errónea división de las oraciones o términos detectados erróneamente por el sistema de reconocimiento de voz-texto.

En la siguiente columna, se muestra el subtítulo traducido automáticamente. Estos segmentos se han obtenido después de descargar los subtítulos automáticos producidos por YouTube. Finalmente, en la última columna, se realiza una posesición manual de los subtítulos.

Además, se muestra una fila sombreada en color rojo claro debajo de cada segmento, en la que se especifican los errores que surgen de cada segmento. Algunos segmentos pueden estar formados por una, dos o tres líneas. Los errores se señalan en diferentes colores según la tipología de error tanto en la tabla como en los gráficos.

Los errores de precisión o *accuracy* se señalan en verde claro. No se han encontrado errores de terminología en los segmentos por lo que no se muestran señalados en ningún color específico. Los errores de *mistranslation*, por otro lado, se señalan en verde y se dividen en diferentes subtipos, como *overly literal*, *false friend*, *should not have been translated*, *entity* y *ASR*. Los errores de *omission* se señalan en rojo, mientras que los de *addition* se señalan en azul claro. Por último, los errores de *untranslated* se señalan en rosa.

Los errores de fluidez o *fluency* se señalan en amarillo claro. Se dividen en diferentes subtipos, como *spelling* que se presenta subrayado, todos los tipos de errores de *typography* están señalados en azul oscuro, los de *grammar* en marrón y *unintelligible* en morado.

Segmento original - Transcripción	Subtítulo traducido automáticamente	Posedición
Sundays are just a nightmare.	Los domingos son como (...)	Los domingos son una pesadilla.
ERRORES	Omission y addition, typography (punctuation)	
It gets to about five o'clock	Si llegaran a las cinco en punto	Llegan las cinco
and songs of praise, I hate,	Y las canciones de alabanza, odio,	y las canciones de alabanza, odio,
absolutely hate songs of praise.	absolutamente odio, las canciones de alabanza	odio profundamente, las canciones de alabanza
ERRORES	Grammar (verbal tenses)	
Comes on and I just think "that's it,	Suenan y creo que eso es todo. Se acabó el fin	suenan y pienso: «eso es todo,
the weekend's over, I've done nothing".	de semana. No he hecho nada	se acabó el fin de semana, no he hecho nada»
ERRORES	Mistranslation (overly literal), typography (Unpaired quote marks)	
And normally I reach for my phone and	y normalmente busco _ teléfono, y	Y normalmente cojo mi teléfono y
text at least five of my friends,	envié un mensaje de texto al menos a cinco de mis amigos	escribo al menos a cinco de mis amigas,
and say, you know, "what are you doing",	y diga que sabe qué está haciendo,	y les digo, ya sabes, «¿qué estás haciendo»
ERRORES	Untranslated, Grammar (verbal tenses), Mistranslation (overly literal), typography (Unpaired quote marks)	
you know. Now being Juliet and one daughter	que ahora es Julia y una hija,	Sabes, ahora siendo Juliet y una hija
you know, we haven't got	que sabe que no tenemos este	sabes, no tenemos eso
this Sunday lunch family thing.	almuerzo familiar del domingo,	de la comida familiar de los domingos.
ERRORES	Mistranslation (overly literal) + Capitalization, Grammar (verbal tenses), Mistranslation (entity), Addition	
So what did you have	entonces, ¿qué tiene el último	Entonces, ¿qué tuvisteis
for last Sunday lunch?	almuerzo del domingo?	el domingo pasado?
ERRORES	Grammar (verbal tenses), Mistranslation (overly literal)	
What did you eat? Tuna sandwich	¿Qué comió atún? sándwich	¿Qué comisteis? Un sándwich de atún
For Sunday lunch?	para algo almuerzo	¿Para comer el domingo?

For Sunday lunch.	para el almuerzo del domingo	Exacto.
ERRORES	Grammar (verbal tenses), typography (punctuation), Capitalization Mistranslation (overly literal)	
From teenager up to mid-20s	desde adolescente hasta mediados de los 20	De adolescente hasta mediados de los 20
I made dinner, a proper dinner.	Hice la cena, Espero que ella cene	preparé la cena, una cena decente
ERRORES	Capitalization, Mistranslation (ASR) and capitalization	
Then I got married to a Master Chef	_Me casé con el maestro chef	Luego me casé con un Master Chef
who took over the kitchen for ten years.	que se hizo cargo de la cocina durante diez años	que se encargó de la cocina durante diez años
ERRORES	Omission, Mistranslation (Should not have been translated & overly literal)	
Really?	(omission)	¿En serio?
And didn't let me in.	y no me dejó entrar	Y no me dejó entrar.
Oh really?	Realmente	¿En serio?
ERRORES	Capitalization, Mistranslation (overly literal)	
And then when we divorced	y luego cuando nos divorciamos	Y cuando nos divorciamos
couldn't remember what to do.	no podía recordar qué hacer,	no podía recordar qué hacer.
ERRORES		
So, your husband, your ex-husband	así que ve a tu esposo, a tu ex esposo	Así que, tu marido, tu exmarido
is to blame for this one. Absolutely.	no le gusta esto,	es el culpable. Desde luego.
ERRORES	Capitalization, spelling, mistranslation (ASR)	
Yeah, one with pickled egg bollocks inside here,	sí, bueno, no pepinillos t +	Sí, uno con los cojones en vinagre aquí dentro
that's it inside your cupboard.	aquí adentro, eso es todo dentro de tu armario,	¿es todo lo que hay en tu armario?
ERRORES	Unintelligible, mistranslation (overly literal), typography (punctuation)	
What's he doing in there?	¿qué está haciendo allí?	¿Qué hace esto aquí?
Ah, what is he doing in there?	ah, ¿qué está haciendo allí?	Ah, ¿qué es lo que está haciendo aquí?
ERRORES	Capitalization, mistranslation (overly literal)	

Well, it was saving me buying a takeaway	bueno, me estaba ahorrando comprar un takeawa	Bueno, me estaba ahorrando comprar
jerk seasonings, ah.	y condimentos idiotas ah,	condimentos jerk para llevar
ERRORES	<u>Capitalization</u> , <u>punctuation</u> , <u>untranslated</u> , <u>mistranslation (Should not have been translated)</u>	
We should put that one next to him.	<u>deberíamos</u> poner ese al lado de él, oh sí	Deberíamos ponerlo a su lado.
There you go. You know,	sabes que	Ahí lo tienes. Ya sabes,
that sits there nice and comfortably.	se sienta allí agradable y cómodamente.	queda bonito y cómodo.
ERRORES	<u>Capitalization</u> , <u>addition</u> , <u>mistranslation (overly literal)</u>	
I like the idea of your friends coming around.	Me gusta la idea de que tus amigos vengan.	Me gusta la idea de que vengan tus amigas.
I like the idea of something	Me gusta la idea de algo	Me gusta la idea de algo
that they can share.	que puedan compartir.	que puedan compartir.
ERRORES	<u>Grammar (word order)</u>	
They can go in the middle of the table	Pueden ir en el medio de la mesa.	Puede ir en el centro de la mesa
and they can all dive in.	y todos pueden sumergirse,	y todas pueden zambullirse.
Yeah, that's perfect.	sí, eso es perfecto,	Sí, es perfecto.
ERRORES	<u>Grammar (agreement)</u>	
Paella, in terms of, you know, usually	Tyler , en términos de que sabes , generalmente	Paella, en términos de que, ya sabes, suele
involve rice, rice and chicken. Yeah.	involucra arroz, arroz y pollo, sí,	involucrar arroz, arroz y pollo. Sí.
ERRORES	<u>Mistranslation (ASR)</u> , <u>Omission</u> , <u>punctuation</u>	
What's wrong with rice?	¿ qué tiene de malo el arroz?	¿Qué tiene de malo el arroz?
Well, it's cooking it,	Bueno, se está cocinando	Bueno, es cocinarlo,
You can make disasters with rice	¿ no puedes hacer desastres_?	puedes hacer desastres con el arroz
ERRORES	<u>Capitalization</u> , <u>Mistranslation (overly literal)</u> , <u>addition</u> , <u>punctuation</u> , <u>omission</u>	
No no no, but I mean, with the stock,	No no, pero con el fondo ,	No, no. Pero con el sofrito,
water, in fact, and things like chorizo	como chorizo y otras cosas	agua, de hecho, y cosas como chorizo
and saffron, it cooks itself.	cocina tú mismo	y azafrán, se cocina solo.
ERRORES	<u>Omission</u> , <u>mistranslation (ASR)</u>	

300 grams of squid and	300 gramos de calamar y	300 gramos de calamar y
300 grams of prawns, please. Thank you.	entrenando los gramos de gambas por favor gracias	300 gramos de gambas, por favor. Gracias.
ERRORES Mistranslation (ASR), punctuation		
What did you have for Sunday lunch last week?	por el almuerzo del domingo la semana pasada	¿Qué comiste el domingo pasado?
Rosted chicken, I think.	en la mesa, mira	Pollo asado, creo.
ERRORES Omission, punctuation, unintelligible		
How many were you at the table Sunday lunch?	(omission)	¿Cuántos estabais en la mesa el domingo?
We're about ten. There you go, see...	(omission)	Eramos unos diez. Ahí tienes, ves...
ERRORES omission		
Paella, when was the last time you ate one?	Tyler fue así como nunca puedes estar	Paella, ¿cuándo comiste una por última vez?
Never. Okay, chorizo. Just smell it.	bien Sarita	Nunca. Vale, chorizo. Solo huélelo.
ERRORES Unintelligible, omission, mistranslation (ASR)		
Nice, spicy and it actually flavours	simplemente no es tan picante y en realidad	Agradable, picante y da sabor
the rice at the same time.	sabores el arroz es casi al mismo tiempo	al arroz al mismo tiempo.
ERRORES Mistranslation (ASR), unintelligible		
Good, get hold of the chilies.	bueno agarra los chiles	Bien, coge los chiles.
Why are you taking the seeds out?	(omission)	¿Por qué quitas las semillas?
ERRORES omission		
Well, they are really hot, and I want to use	están muy picantes y quiero usar	Bueno, son muy picantes, y quiero utilizar,
the sort of whole chili	este tipo de chile entero	más o menos, el chile entero,
rather than the seeds.	en lugar de las semillas	en lugar de las semillas.
ERRORES Capitalization, mistranslation (false friend), mistranslation (overly literal), punctuation		
Slice it down like that. And a nice thing about	cortarlo así es bueno comenzar el árbol el estás	Córtalo así. Y lo bueno de
sweating the chorizo down with the onions...	abajo en las cebollas	sofreir el chorizo con las cebollas...
Oh, what was that? The chili	(omission)	¡Oh! ¿Qué pasa? El chili...
ERRORES Unintelligible, omission		
You see, that nice romantic moment	fue ese lindo momento romántico	Lo ves... Un bonito momento romántico

when you're not crying over a phone bill, but you're crying over a paella.	cuando no estás llorando por una factura telefónica tu cliente en Pilar	cuando no lloras por una factura del teléfono pero estas llorando por una paella
ERRORES Capitalization, Mistranslation (overly literal), unintelligible		
I'm using chicken but you also make this with fish or rabbit.	está usando pollo pero también lo haces con pescado o conejo	Estoy usando pollo, pero también puedes hacerla con pescado o conejo.
Adding sherry gives it a deeper flavour.	Agregando jerez le da un sabor más profundo lo	Añadir jerez le da un sabor más intenso.
ERRORES Capitalization, grammar (agreement), mistranslation (overly literal), addition, punctuation		
Once that's evaporated, what I want to do is pour the rice in.	que evaporó lo Lo que quiero que hagas es poner el arroz	Una vez que se haya evaporado, lo que hacemos es echar el arroz.
ERRORES Unintelligible, addition, mistranslation (overly literal)		
And this is where you always get a little bit confused, isn't it?	y aquí es donde siempre te confundes un poco	Y aquí es donde siempre nos confundimos un poco, ¿no?
Sprinkle the paella, mix it all in.	espolvorea la mezcla Pilar en nadie	Espolvorea la paella, mézclalo todo.
ERRORES Capitalization, punctuation (question marks), mistranslation (entity), unintelligible		
I want to get that rice nicely coloured. Okay. There you go.	para que el arroz tenga un color agradable, está bien, ahí tienes,	Quiero que ese arroz tenga un buen color. Vale, lo tenemos.
ERRORES Capitalization, mistranslation (overly literal)		
So next I want you to do is to get the saffron.	así que la próxima vez lo que haces es conseguir el azafrán	Lo siguiente que quiero que hagas es coger el azafrán.
ERRORES Capitalization, mistranslation (overly literal), punctuation		
What does this do? The saffron gives it an amazing colour and gives it a really nice flavour as well.	¿Qué hace esto? El azafrán le da un color increíble y también le da un sabor muy agradable.	¿Para qué sirve? El azafrán le da un color increíble y también un sabor muy agradable.
And to finish off prawns, squid, and clams.	Y para terminar, gambas, calamares y almejas.	Y para terminar gambas, calamares y almejas.
ERRORES Mistranslation (overly literal)		
So far, so good. Good, you're doing good.	(omission) (omission)	Hasta aquí, todo bien. Bien, lo estás haciendo bien.

A bit of pepper there at the back	(omission)	Un poco de regusto a pimienta.
ERRORES	omission	
These tomatoes will bring that down.	baje eso y ahí	Los tomates lo rebajarán.
And there we are. Look, paella, fit for a king.	estamos, parecemos bastante dignos de un rey	Y ya está. Mira, una paella digna de un rey.
And then, my dear lady, we're ready to roll.	Y luego, mi querida señora, estamos listos para rodar	Así, mi querida, estamos listos para la acción.
ERRORES	Unintelligible, grammar (agreement), mistranslation (overly literal)	
Ladies! Oh, yes!	señoras	¡Damas! ¡Oh, sí!
It's been a long time since	veces mientras me sentaba	Hacía mucho tiempo desde que
I sat around a table like this, ladies.	alrededor de una mesa como esta.	me sentaba en una mesa así, queridas.
ERRORES	Omission, capitalization, omission	
I'm starting to feel slightly intimidated.	Estoy un poco intimidado.	Empiezo a sentirme un poco intimidado.
I'm gonna ban texting,	Voy a prohibir los mensajes de texto, ¿de acuerdo?	Voy a prohibir los mensajes,
no more texts on a Sunday afternoon.	no más Texas bajo algo por la tarde	no más mensajes los domingos por la tarde.
ERRORES	Addition, mistranslation, unintelligible	
No, because you don't need to text.	Porque estarás aquí	No, porque no necesitáis enviar mensajes.
You don't need to be that frustrated	Dispositivo	No tenéis que frustraros tanto
because you're all going to be here.	(omission)	porque vais a estar todas aquí.
ERRORES	unintelligible	
It's like a resort rice.	(omission)	Es como un arroz de resort.
However, if you can't find paella rice,	sin embargo, si no puede encontrar el	Aunque si no encuentras arroz de paella,
then choose the long grain.	archivo, entonces elija el gris largo.	elige el arroz largo.
ERRORES	Capitalization, mistranslation (ASR), grammar (agreement)	
Madam, you have to promise me something today.	Tiene que prometerme algo hoy.	Querida, tienes que prometerme algo hoy.
That once a week, you will start cooking	Sí, eso una vez a la semana. Sí. Comenzará a	Que una vez a la semana vas a empezar a
something from scratch. From scratch.	cocinar algo desde cero	cocinar algo desde cero. Desde cero.
ERRORES	grammar (agreement), mistranslation (overly literal)	
Good girl. Make a decision to go shopping	(omission)	Buena chica. He decidido ir a hacer la compra
with my list, I'm going to moat around the	Voy a dar vueltas por ese supermercado	con mi lista, voy a recorrer

supermarket avoiding all the	evitando todas las destrucciones precocinadas	el supermercado evitando toda la comida
ERRORES	Omission, mistranslation (overly literal)	
pre-packed, precooked, pre-ready-made.	y listas para usar y	preenvasada, precocinada, preelaborada.
And I'm definitely going to cook.	definitivamente voy a cocinar	Y, de una vez por todas, voy a cocinar.
ERRORES	Mistranslation (overly literal)	

Tabla 2. Tabla de análisis de los errores de traducción: segmentos originales, traducción automática y posesión.

Una vez detectados y señalados los errores, procedemos a clasificarlos en función de su tipología.

4.2. Resultados por tipología de error

Tras haber identificado los errores, procedemos a clasificarlos según hagan referencia a la fluidez y precisión empleando gráficos. Cada gráfico muestra el porcentaje de errores acompañado de una leyenda.

Según lo obtenido, el 54 % de los errores pueden clasificarse como precisión (*accuracy*) y el 46 % como fluidez (*fluency*).

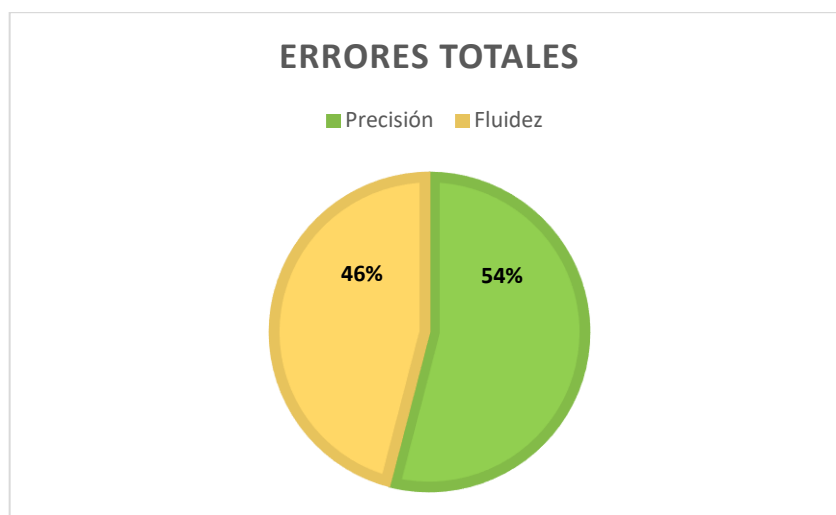


Gráfico 1. Tipología de errores

4.2.1. Errores de precisión

Los errores de precisión se dividen en varias categorías según Ortiz (2016: 63), siendo el error de traducción incorrecta el más común, que representa el 61 % del total de los errores de precisión. Esto se debe a que las traducciones automáticas frecuentemente se basan en el análisis de patrones lingüísticos y estadísticas para producir una traducción. Sin embargo, estas técnicas no siempre tienen en cuenta el contexto de la oración o del texto completo, lo que puede llevar a traducciones demasiado literales que no reflejan el significado completo de la oración original.

Dentro de este error de traducción, los errores de traducción demasiado literales (*mistranslation, overly literal*) representan el 65 % de los casos, lo que indica que el sistema de traducción automática comprende las palabras literalmente en lugar de entender el significado completo de la oración y el contexto general de la muestra. Ejemplos de este tipo de error son «*what's he doing in there?*», subtítulo automáticamente como «¿qué está haciendo allí?» y que hemos poseído como «¿Qué hace esto aquí?» o «*that sits there nice and comfortably*» que ha sido traducido demasiado literal como «se sienta allí agradable y cómodamente» y cuya traducción adecuada sería «queda bonito y cómodo».

Los errores de falsos amigos (*mistranslation, false friend*) representan solo el 2 %, lo que sugiere que el modelo está lidiando efectivamente con el significado de las palabras y que el contenido del vídeo que se está analizando no ofrece mayor dificultad respecto a este parámetro. Lo vemos reflejado en «*Well, they are really hot, and I want to use*», que el sistema de traducción automática ha traducido como «están muy picantes y quiero usar», en vez de utilizar el verbo «ser» que es más adecuado al contexto, «Bueno, son muy picantes, y quiero utilizar.».

Asimismo, hemos notado una tendencia a traducciones incorrectas debidas a que YouTube no es capaz de captar correctamente la oración a través del reconocimiento de voz-texto. A este error específico de *mistranslation* lo hemos denominado ASR.

El 23 % de los errores de traducción son resultado del sistema de reconocimiento automático de voz (ASR), es decir, el propio sistema ASR de YouTube no ha reconocido correctamente la voz del vídeo y, por lo tanto, ha realizado una traducción errónea del segmento. Como ejemplos de este tipo de errores encontramos «*Okay, chorizo.*» que el sistema de reconocimiento de voz-texto ha interpretado como «bien Sarita» cuando la traducción apropiada sería «Vale, chorizo». O lo encontramos en el segmento «*a proper dinner.*». El sistema de reconocimiento de voz-texto ha interpretado «*I hope her dinner*» (oración incorrecta gramaticalmente) y lo ha traducido automáticamente como «Espero que ella cene».

El 5 % de los errores de traducción se deben a entidades que no se traducen o no deberían traducirse (*mistranslation, should not have been translated*), como «*Then I got married to a Master Chef*» que ha sido traducido automáticamente como «Me casé con el maestro chef» cuando debería haberse mantenido el nombre «Master Chef»; y el 5 % son errores en la traducción de entidades específicas o nombres propios tanto de personas como de platos de comida (*mistranslation, entity*). Por ejemplo, en el segmento «*you know, now being Juliet and one daughter*» se ha traducido el nombre propio como «que ahora es Julia y una hija.».

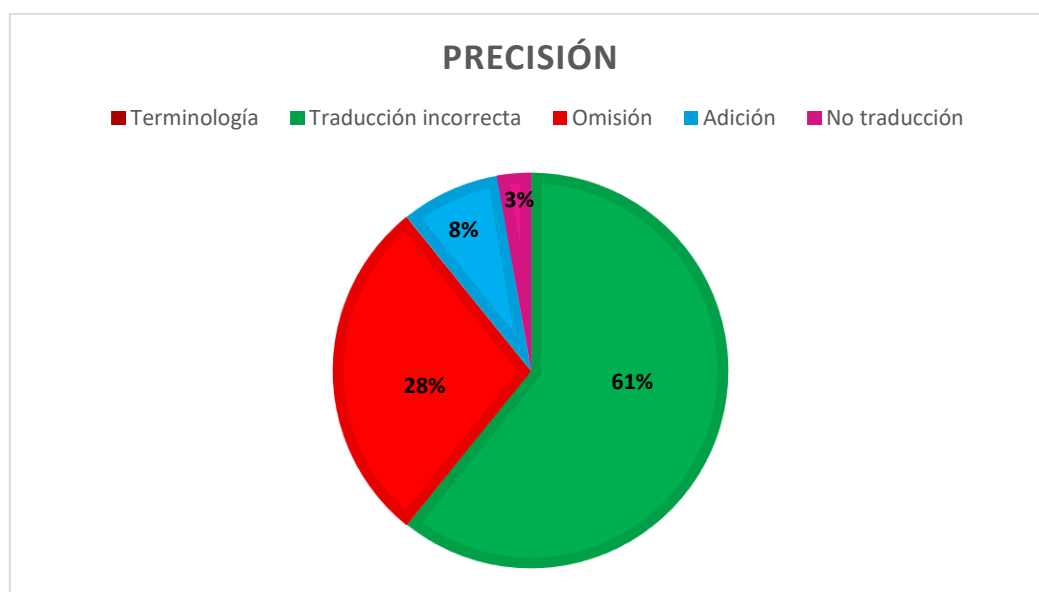


Gráfico 2. Clasificación de los errores de precisión

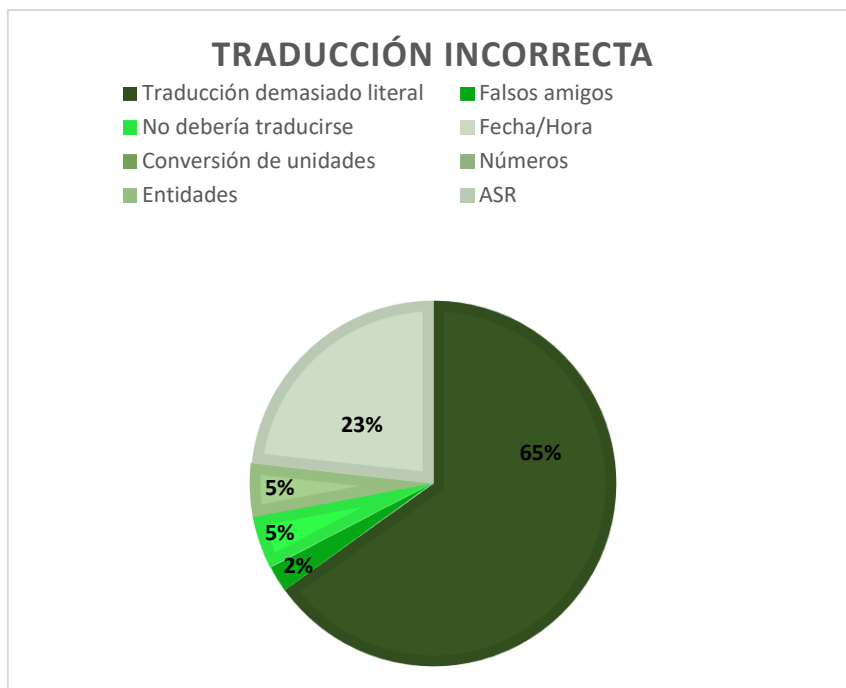


Gráfico 3. Clasificación de las traducciones incorrectas

A continuación, hemos analizado los errores de fluidez.

4.2.2. Errores de fluidez

Los errores de fluidez son aquellos que afectan la cohesión y la coherencia del texto, impidiendo que se transmita el mensaje de manera clara y efectiva. Estos pueden manifestarse de diversas maneras, tales como errores de ortografía (*spelling*), errores de tipografía (*typography*), errores gramaticales (*grammar*) y frases incomprensibles (*unintelligible*).

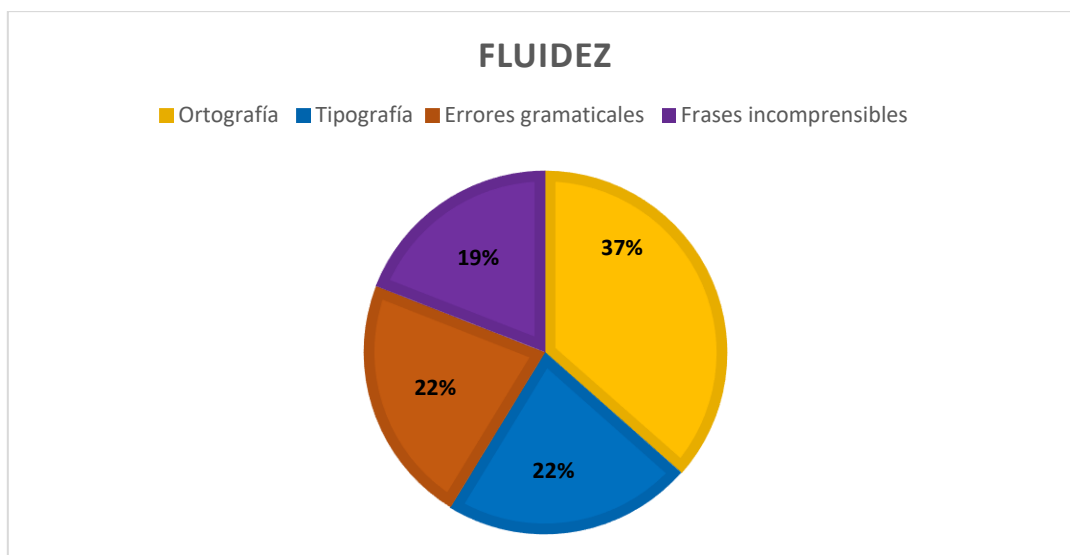


Gráfico 4. Clasificación de los errores de fluidez.

Dentro de esta categoría los errores de ortografía (*spelling*) son el tipo más común y representan el 37 % del total. En especial, los relativos al uso de mayúsculas y minúsculas que ocupan casi el 100 % de los errores de esta categoría, con ejemplos como «*From teenager up to mid-20s / I made dinner, a proper dinner*», que el sistema de traducción automática ha traducido como «desde adolescente hasta mediados de los 20 / Hice la cena, Espero que ella

cene»; aunque también reconocemos un error de ortografía en el uso de prefijos («ex esposo» / «exmarido»).

Los errores de tipografía (*typography*) representan el 22 %, siendo los errores de puntuación los más comunes, es decir, el uso de puntos, comas, signos de interrogación y exclamación representa el 86 % frente al 14 % de los errores de comillas o corchetes no emparejados. Este se comete cuando se utiliza una comilla o un corchete de apertura, pero no se cierra correctamente con su signo correspondiente. Un ejemplo de error tipográfico es «300 grams of prawns, please. Thank you.», que se traduce automáticamente como «entrenando los gramos de gambas por favor gracias». Como error de comillas o corchetes no emparejados encontramos «Comes on and I just think “that’s it, the weekend’s over, I’ve done nothing”», segmento en el que la traducción automática ha omitido las comillas.

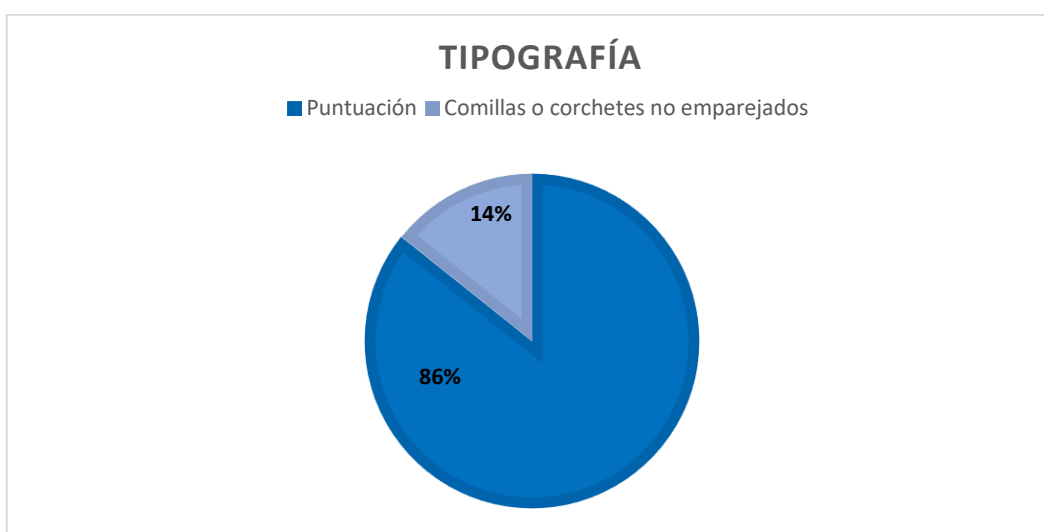


Gráfico 5. Clasificación de los errores de tipografía.

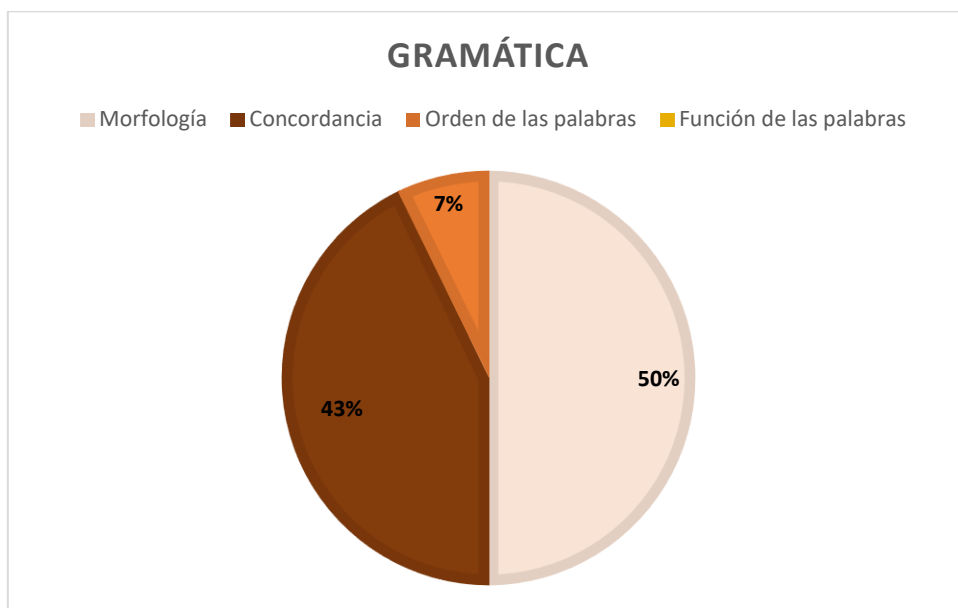
Los errores de gramática también representan el 22 % de los errores de fluidez, con los errores de morfología, es decir, mal uso del morfema asociado a los tiempos verbales (50 %) y de concordancia (*agreement*, 43 %) siendo las principales causas. Por último, el 19 % de los errores de fluidez se deben a problemas de inteligibilidad (*unintelligible*), lo que indica que el modelo no está produciendo un resultado coherente y comprensible por el receptor.

Como error de morfología en el uso de tiempos verbales observamos este ejemplo «*It gets to about five o'clock*» en el que su traducción automática utiliza el condicional «Si llegaran a las cinco en punto» cuando debería utilizarse el presente «Llegan las cinco».

Respecto a la concordancia (*agreement*) encontramos «*I'm using chicken but you also*» que se ha traducido como «está usando pollo pero también», y el verbo no concuerda con el sujeto de la oración.

Otros errores relativos a la inteligibilidad (*unintelligible*) los apreciamos, por ejemplo, en «*Sprinkle the paella, mix it all in.*», que ha sido traducido como «espolvorea la mezcla Pilar en nadie»; «*Slice it down like that. And a nice thing about sweating the chorizo down with the onions...*» cuya traducción es «cortarlo así es bueno comenzar el árbol el estás abajo en las cebollas», entre otros errores señalados previamente en la Tabla 2. Es necesario considerar estos errores y realizar una adecuada posesición, ya que la comprensión del mensaje se ve perjudicada.

Gráfico 6. Clasificación de errores gramaticales.



Después de realizar un análisis exhaustivo y una comparación detallada de los resultados obtenidos, hemos extraído las conclusiones que presentamos en el siguiente capítulo del trabajo.

5. CONCLUSIONES

Durante el desarrollo de este Trabajo de Fin de Grado, hemos tenido la oportunidad de aplicar y poner en práctica diversas competencias y habilidades que nos han permitido alcanzar los objetivos previstos. La realización de este proyecto ha requerido la comprensión de los fundamentos teóricos de la literatura previa, el desarrollo de una metodología de análisis precisa, la extracción y análisis de datos a partir de una muestra, y la capacidad de sintetizar y presentar la información de manera clara y precisa.

Después de realizar nuestro análisis, hemos identificado distintas tipologías de errores que, en esencia, determinan la precisión y la fluidez del texto resultante. El análisis que hemos llevado a cabo muestra que la mayoría de los errores cometidos por los sistemas de traducción automática se relacionan con la precisión, en concreto, con traducciones erróneas y omisiones. Gran parte de estos errores se deben a las limitaciones propias del sistema de reconocimiento de voz-texto y pueden ocurrir por diversas razones, tales como la calidad del audio, el ruido de fondo, la complejidad del lenguaje, las variaciones de los acentos o la capacidad limitada del modelo de reconocimiento. Los sistemas de traducción automática tienden a la traducción literal, es decir, a traducir palabra por palabra, sin tener en cuenta el contexto o el sentido global de la frase. Esto se debe a que los modelos de TA están basados en algoritmos que buscan patrones y correlaciones en grandes conjuntos de datos lingüísticos, y no tienen una comprensión completa del significado y la intención detrás del texto que están traduciendo.

Entre los errores de fluidez destacan, en particular, los errores de tipografía, como la puntuación incorrecta, y la inteligibilidad o utilización de las mayúsculas, que son comunes en estos sistemas. Estos errores no solo afectan la apariencia del texto, sino que también pueden dificultar la comprensión de este, lo que puede obstaculizar la comunicación efectiva entre los usuarios de diferentes idiomas. Aunque no todos los segmentos se ven afectados por estos errores, es importante tener en cuenta que incluso un pequeño número de errores puede tener un impacto significativo en la calidad general del texto y la comprensión del mensaje.

Por otro lado, nuestro análisis ha revelado una serie de limitaciones y errores en los sistemas de reconocimiento de voz-texto y de traducción automática. Creemos que tienen un camino largo de mejora por delante, especialmente en el caso de plataformas como YouTube. Estos sistemas pueden resultar de gran ayuda para tareas como la transcripción de audio o la creación de subtítulos, aunque se requiere una posesición manual que tenga en cuenta las restricciones propias de los subtítulos para garantizar la calidad del producto final, especialmente en el caso de documentos audiovisuales y subtítulos, donde la precisión y la coherencia son esenciales.

Los sistemas de traducción automática se aplican en muchos campos, entre los que se incluye la traducción de documentos oficiales, de textos en línea, de subtítulos para películas y series, de textos técnicos en campos como la medicina o la informática, etc. Los sistemas de reconocimiento de voz-texto son de gran ayuda para realizar transcripciones de conferencias y su posterior traducción y subtitulación, y pueden ser especialmente útiles a la hora traducir grandes cantidades de textos o documentos audiovisuales de manera rápida y eficiente.

Por último, la metodología utilizada en este estudio podría aplicarse a otras disciplinas y lenguas. Además, se trata de una tecnología en constante evolución y desarrollo, por lo que se podría replicar en un futuro y obtener resultados completamente diferentes que podrían contrastarse con los presentados en este TFG para valorar la evolución de los sistemas de traducción automática y de reconocimiento de voz-texto.

6. BIBLIOGRAFÍA

- Allen, J. (2003). Post-editing. En H. Somers (Ed.), *Computers and translation: a translator's guide*. Ámsterdam / Filadelfia: John Benjamins Publishing, pp. 297-318. Disponible en https://www.academia.edu/63151678/Post_editing
- Ávila Cornejo, E. (2016). *Sistema intercomunicador mediante reconocimiento de voz y texto a voz utilizando Alexa y Raspberry Pi*. Reporte Técnico de Investigación. Ciudad Juárez: Universidad Autónoma de Ciudad Juárez. Disponible en <http://erecursos.uacj.mx/handle/20.500.11961/3000>
- Ávila-Moreno, C. D., Herrán-Camargo, K. S., Muñoz-Bernal, A. C., Rodríguez-Rojas, N., & Torres-Ortiz, D. C. (2020). La subtítulos audiovisual en un mundo globalizado, ¿cumple su función? *Comunicación, cultura y política*, 9, 39-46. DOI: [10.21158/21451494.v9.n0.2018.2729](https://doi.org/10.21158/21451494.v9.n0.2018.2729)
- Casacuberta Nolla, F.; Peris Abril, A. (2017). Traducción automática neuronal. *Revista Tradumàtica. Tecnologies de la Traducció*, 15, 66-74. DOI: [10.5565/rev/tradumatica.203](https://doi.org/10.5565/rev/tradumatica.203)
- Cid-Leal, P., Espín-García, M., & Presas, M. (2019). Traducción automática y posesión: perfiles y competencias en los programas de formación de traductores. *MonTI - Monografías de Traducción e Interpretación*, 11, 187-214. DOI: [10.6035/monti.2019.11.7](https://doi.org/10.6035/monti.2019.11.7)
- Díaz Cintas, J. (2003). *Teoría y práctica de la subtítulos inglés-español*. Barcelona. Editorial Ariel.
- Díaz, A. P., Suarez, H. J. B., Lara-Díaz, M. F., Moreno, A. M., & Díaz, M. J. S. (2022). Una aproximación a la lectura de subtítulos, percepción y comprensión de piezas audiovisuales mediante tecnología de seguimiento ocular. *Revista iberoamericana de psicología*, 15(2), 13-23. DOI: [10.33881/2027-1786.rip.15202](https://doi.org/10.33881/2027-1786.rip.15202)
- Durán, I., & Del Moral Álvarez, J. (2014). Competencia documental para la traducción agroalimentaria EN-ES: fuentes de información y su evaluación. *SKOPOS. Revista Internacional de Traducción e Interpretación*. DOI: [10.21071/skopos.v5i.4288](https://doi.org/10.21071/skopos.v5i.4288)
- Epstein, B. J. (2018). What's Cooking: Translating Food. *Translation Journal*, 13(3). Disponible en <http://www.translationjournal.net/journal/49cooking.htm>
- Forcada, M.L., Sánchez, F. y Pérez, J.A. (2016). *Manual de informática y de tecnologías para la traducción*. Alicante: Departamento de Lenguajes y Sistemas Informáticos de la Universidad de Alicante. Disponible en <http://hdl.handle.net/10045/53085>
- García, J. (2021, 18 marzo). *Los subtítulos automáticos en tiempo real de Google Chrome ya están disponibles: así funcionan y se activan*. Xataka. Disponible en <https://www.xataka.com/aplicaciones/subtitulos-automaticos-tiempo-real-google-chrome-estan-disponibles-asi-funcionan-se-activan>
- Kaushik, L., Sangwan, A., & Hansen, J. H. L. (2013). Automatic sentiment extraction from YouTube vídeos. *IEEE Automatic Speech Recognition and Understanding Workshop*. DOI: [10.1109/asru.2013.6707736](https://doi.org/10.1109/asru.2013.6707736)
- Lommel, A., Uszkoreit, H., & Burchardt, A. (2014). Multidimensional Quality Metrics (MQM): A Framework for Declaring and Describing Translation Quality Metrics. *Revista Tradumàtica, Tecnologies de la Traducció*, 12, 455. DOI: [10.5565/rev/tradumatica.77](https://doi.org/10.5565/rev/tradumatica.77)
- Malik, M., Malik, M. I., Mehmood, K., & Makhdoom, I. (2020). Automatic speech recognition: a survey. *Multimedia Tools and Applications*, 80(6), 9411-9457. DOI: [10.1007/s11042-020-10073-7](https://doi.org/10.1007/s11042-020-10073-7)

- Maldonado, M. C. y Liébana, M. (2021). Los motores de traducción automática y su uso como herramienta lexicográfica en la traducción de unidades léxicas aisladas. *Círculo de Lingüística Aplicada a la Comunicación*, 88: 189-212. Disponible en <https://doi.org/10.5209/clac.77002>
- Metola Navaridas, M. (2022). *La traducción automática neuronal inglés-español de las recetas de cocina: análisis de errores*. TFG. Valladolid: Universidad de Valladolid. Disponible en <https://uvadoc.uva.es/handle/10324/54877>
- O'Brien, S. (2011). Towards predicting post-editing productivity. *Machine Translation*, 25(1): 195-225. DOI: [10.1007/s10590-011-9096-7](https://doi.org/10.1007/s10590-011-9096-7)
- Ortiz, C. (2016). *Implementing Machine Translation and Post-Editing to the Translation of Wildlife Documentaries through Voice-over and Off-screen Dubbing*. Tesis doctoral. Barcelona: Universitat Autònoma de Barcelona. Disponible en <http://hdl.handle.net/10803/400020>
- Rebollo Trigueros, J. (2018). *El subtítulo palabra por palabra: Aplicación de los sistemas de lectura rápida RSVP y ORP en los subtítulos para soportes de tamaño reducido*. TFG. Granada: Universidad de Granada. Disponible en <https://digibug.ugr.es/handle/10481/51585>
- Rendueles, C., & Sádaba, I. (2019). Digitalización y cambio social. De las expectativas apocalípticas a la tecnopolítica del presente. *Cuadernos de Relaciones Laborales*, 37(2), 331-349. DOI: [10.5209/crla.66041](https://doi.org/10.5209/crla.66041)
- Ruiz Romero, M. D. L. Á. (2020). *Tipología textual y traducción en el ámbito agroalimentario: definición del perfil y ejercicio profesional*. Tesis Doctoral. Córdoba: Universidad de Córdoba. Disponible en <https://helvia.uco.es/xmlui/bitstream/handle/10396/20185/2020000002097.pdf?sequence=3&isAllowed=y>
- Szarkowska, A., & Gerber-Morón, O. (2018). Viewers can keep up with fast subtitles: Evidence from eye movements. *PLOS ONE*, 13(6). DOI: [10.1371/journal.pone.0199331](https://doi.org/10.1371/journal.pone.0199331)
- Sánchez Ramos, M. M., Rico Pérez, C. (2020). *Traducción automática. Conceptos clave, procesos de evaluación y técnicas de posesición*. Granada: Comares.
- Sedeño Valdellós, A. M. (2011). Cine y medios audiovisuales ante la globalización. *Encuentros*, 9(1), 11-20. Disponible en <https://dialnet.unirioja.es/descarga/articulo/4049574.pdf>
- Ubillús Sánchez, C. K. (2022). *El uso de la traducción automática de Google Cloud en un proceso de posesición*. TFG. Santiago de Surco: Universidad Peruana de Ciencias Aplicadas. Disponible en <https://repositorioacademico.upc.edu.pe/handle/10757/661194>
- Veale, T. y Way, A. (1997). Gaijin: a bootstrapping approach to example-based machine translation. *Proceedings of the 2nd international conference, recent advances in natural language processing*, pp. 239-244. Disponible en: <https://aclanthology.org/www.mt-archive.info/RANLP-1997-Veale.pdf>
- Wagner, E. (1985). Post-editing systran: a challenge for commission translators. *Terminologie et traduction*, 3, 1-7. Disponible en https://www.academia.edu/63151678/Post_editing